THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 17, pp. 11832–11840, April 25, 2008

Molecular Insights into the Biosynthesis of the F_{420} Coenzyme*

Received for publication, December 19, 2007, and in revised form, January 28, 2008 Published, JBC Papers in Press, February 5, 2008, DOI 10.1074/jbc.M710352200

Farhad Forouhar[‡], Mariam Abashidze[‡], Huimin Xu[§], Laura L. Grochowski[§], Jayaraman Seetharaman[‡], Munif Hussain[‡], Alexandre Kuzin[‡], Yang Chen[‡], Weihong Zhou[‡1], Rong Xiao[¶], Thomas B. Acton[¶], Gaetano T. Montelione[¶], Anne Galinier[‖], Robert H. White[§], and Liang Tong[‡2]

From the [‡]Department of Biological Sciences, Northeast Structural Genomics Consortium, Columbia University, New York, New York 10027, the [§]Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, [¶]Center for Advanced Biotechnology and Medicine, Northeast Structural Genomics Consortium, Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, New Jersey 08854, and [‖]Laboratoire de Chimie Bactérienne, UPR 9043, Institut de Biologie Structurale et Microbiologie, CNRS, 31 Chemin Joseph Aiguier, 13009 Marseille, France

Coenzyme F_{420}, a hydride carrier, is found in Archaea and some bacteria and has crucial roles in methanogenesis, antibiotic biosynthesis, DNA repair, and activation of antitubercular compounds. CofD, 2-phospho-L-lactate transferase, catalyzes the last step in the biosynthesis of F_{420}-0 (F_{420} without polyglutamate), by transferring the lactyl phosphate moiety of lactyl(2)diphospho-(5′)guanosine to 7,8-didemethyl-8-hydroxy-5-deazariboflavin ribitol (Fo). CofD is highly conserved among F_{420}-producing organisms, and weak sequence homologs are also found in non-F_{420}-producing organisms. This superfamily does not share any recognizable sequence conservation with other proteins. Here we report the first crystal structures of CofD, the free enzyme and two ternary complexes, with Fo and P_i or with Fo and GDP, from *Methanosarcina mazei*. The active site is located at the C-terminal end of a Rossmann fold core, and three large insertions make significant contributions to the active site and dimer formation. The observed binding modes of Fo and GDP can explain known biochemical properties of CofD and are also supported by our binding assays. The structures provide significant molecular insights into the biosynthesis of the F_{420} coenzyme. Large structural differences in the active site region of the non-F_{420}-producing CofD homologs suggest that they catalyze a different biochemical reaction.

The coenzyme F_{420} is a hydride carrier that is found in Archaea and in high G+C Gram-positive bacteria such as *Streptomyces* and *Mycobacterium*. This coenzyme is essential for energy metabolism in methanogenic Archaea, which involves the conversion of CO_2, several other one-carbon compounds, and acetate to CH_4 (1–4). In *Streptomyces*, F_{420} is involved in the biosynthesis of tetracycline, lincomycin, and other natural products (5–9). In some cyanobacteria, a precursor of F_{420} is a cofactor in DNA photolyases for DNA repair (10, 11). In *Mycobacterium tuberculosis*, an F_{420}-dependent glucose-6-phosphate dehydrogenase is required for the reductive activation of a series of nitroimidazo-oxazine compounds for their antitubercular effects (12–14).

Coenzyme F_{420} is named for the intense absorption at 420 nm by the oxidized form of this compound. The chromophore is 7,8-didemethyl-8-hydroxy-5-deazariboflavin, which is linked to a ribityl group at its N-10 position to produce Fo[3] (Fig. 1*A*). Fo is covalently linked to a lactyl phosphate (LP) group to produce F_{420}-0 (Fig. 1*A*). Finally, mature F_{420} coenzymes in different organisms contain 1–7 glutamate residues (F_{420}-*n*) that are covalently linked by amide bonds through their α- and/or γ-carboxylates (1, 15–17). Although F_{420} contains a (deaza)riboflavin moiety, its biochemistry is actually more similar to that of $NAD(P)^+$ than FMN/FAD. However, F_{420} has a lower midpoint potential (−360 mV) than $NAD(P)^+$ (−320 mV) (13).

The biosynthesis of F_{420} uses lactate, 4-hydroxyphenylpyruvate (an intermediate in tyrosine biosynthesis), and a pyrimidine intermediate in riboflavin biosynthesis as precursors (1, 18–21). The last step in the biosynthesis of F_{420}-0 is the condensation of Fo and lactyl(2)diphospho-(5′)guanosine (LPPG), catalyzed by the enzyme 2-phospho-L-lactate transferase (CofD) (Fig. 1*A*) (20, 21). Lactyl(2)diphospho(5′)adenosine can also be used as the donor of the LP group, but the enzyme is much less efficient with this substrate (30-fold higher K_m and 14-fold lower k_{cat} values) (20). The catalysis by CofD requires Mg^{2+} ions, and the enzyme can also catalyze the phosphorylation of Fo using GTP, GDP, PP_i, or PPP_i as the donor (20).

The amino acid sequence of CofD is well conserved among archaeal and bacterial organisms that are known to produce F_{420} (Fig. 1*B*). Weak sequence homologs (~20% identity) are also found in some bacteria (such as *Bacillus* and *Escherichia coli*) that do not produce F_{420} (Fig. 1*B*), but the biochemical functions of these homologs are not known (22). CofD is a member of pfam family UPF0052. There are 381 sequences in

* This work was supported by Protein Structure Initiative of the National Institutes of Health Grants P50 GM062413 and U54 GM074958, National Science Foundation Grant MCB 0231319, the CNRS, and the French Foundation for Medical Research. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 3CGW, 3C3D, and 3C3E) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Present address: College of Life Sciences, Nankai University, Tianjin, 300071, China.

[2] To whom correspondence should be addressed. E-mail: ltong@columbia.edu.

[3] The abbreviations used are: Fo, 7,8-didemethyl-8-hydroxy-5-deazariboflavin ribitol; LPPG, lactyl (2) diphospho-(5′)guanosine; mant, *N*-methylanthraniloyl; PDB, Protein Data Bank; LP, lactyl phosphate; TES, 2-{[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]amino}ethanesulfonic acid.

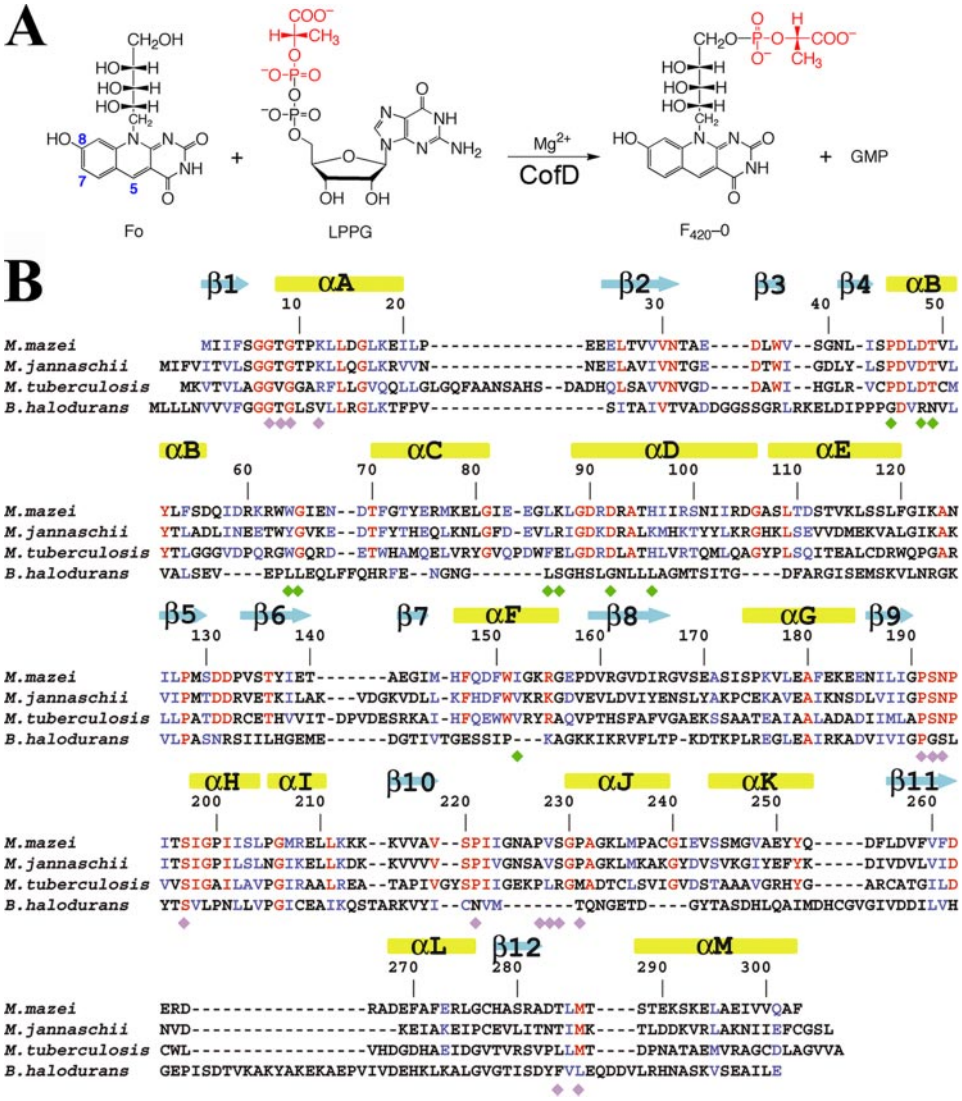

FIGURE 1. **Sequence alignment of CofD homologs.** *A,* chemical reaction catalyzed by CofD. *B,* sequence alignment of CofD from *M. mazei*, *M. jannaschii*, and *M. tuberculosis*, as well as the BH3568 protein from *B. halodurans*. The secondary structure elements are labeled. Residues involved in GDP binding are indicated with the *purple diamonds* and those involved in Fo binding with *green diamonds*.

this pfam family, including eukaryotic homologs. Besides these, the CofD enzymes do not share any recognizable homology with other proteins in the sequence data base (20), and there is no structural information on CofD. We report here the first crystal structures of the CofD enzyme from *Methanosarcina mazei*, in its free form at 3.1 Å resolution, in a ternary complex with Fo and P_i at 2.5 Å resolution, and in a ternary complex with Fo and GDP at 3.0 Å resolution.

MATERIALS AND METHODS

Protein Expression and Purification—The production of CofD protein was carried out as part of the high throughput protein production process of the Northeast Structural Genomics Consortium (23). CofD corresponds to Northeast Structural Genomics Consortium Target MaR46. The full-length *cofD* gene from *M. mazei* (strain *Goe1*) was cloned into a pET21d (Novagen) derivative, generating plasmid pMaR46-21. The resulting recombinant protein contains eight non-native residues (LEHHHHHH) at the C terminus. The construct was verified by standard DNA sequence analysis. *Escherichia coli* BL21 (DE3) pMGK cells, a rare codon-enhanced strain, were transformed with pMaR46-21. A single isolate was cultured in MJ9 minimal media (24) supplemented with selenomethionine, lysine, phenylalanine, threonine, isoleucine, leucine, and valine for the production of selenomethionine-labeled CofD (25). Initial growth was carried out at 37 °C until the A_{600} of the culture reached 0.6–0.8. The incubation temperature was then decreased to 17 °C, and protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside at a final concentration of 1 mM. Following overnight incubation, the cells were harvested by centrifugation.

Selenomethionyl CofD was purified by standard methods. Cell pellets were resuspended in lysis buffer (50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 mM imidazole, and 5 mM β-mercaptoethanol) and disrupted by sonication. The resulting lysate was clarified by centrifugation at 26,000 × *g* for 45 min at 4 °C. The supernatant was loaded onto a nickel-nitrilotriacetic acid column (Qiagen) and eluted in lysis buffer containing 250 mM imidazole. Fractions containing partially purified CofD were pooled, and buffer conditions providing monomeric samples were optimized by analytical gel filtration detected by static light scattering, following the protocol described elsewhere (23). Preparative gel filtration (Superdex 75, GE Healthcare) was then performed using a buffer containing 10 mM Tris (pH 7.5) and 5 mM dithiothreitol. The purified CofD protein was concentrated to 10 mg/ml, flash-frozen in aliquots, and used for crystallization screening. Sample purity (>97%) and molecular weight (34.77 kDa) were verified by SDS-PAGE and matrix-assisted laser desorption ionization time-of-flight mass spectrometry, respectively. The yield of purified protein was ~50 mg/liter of culture.

Cloning, expression, and purification of MJ1256 from *Methanocaldococcus jannaschii* were performed as described previously (20). The D50A mutant of MJ1256 was created with the QuikChange site-directed mutagenesis kit (Stratagene), according to the manufacturer's protocol. The mutated plasmid was sequenced to verify the incorporation of the correct mutation.

Protein Crystallization—The CofD free enzyme and two ternary complexes (with Fo + P_i and Fo + GDP) were crystallized at 20 °C by the hanging-drop vapor diffusion method. For the free enzyme crystals, 2 μl of protein solution containing CofD (10 mg/ml in 5 mM Tris (pH 7.5), 100 mM NaCl, and 5 mM dithiothreitol) were mixed with 2 μl of the reservoir solution consisting of 16% (w/v) PEG3350 and 200 mM $LiNO_3$. Crystals of CofD in complex with Fo and P_i were grown by mixing 2 μl of the protein solution containing 5 mM Fo and 5 mM LP (only phosphate is observed in the final structure) with 2 μl of 4.2 M sodium formate, whereas the crystals of CofD in complex with Fo and GDP were grown by mixing 2 μl of protein solution containing 5 mM Fo, 5 mM GTP (but only GDP is observed in the final structure), and 20 mM magnesium chloride with 2 μl of reservoir solution of 100 mM Hepes (pH 7.15), 1.35 M ammonium sulfate, and 0.5% (w/v) PEG8000.

The free enzyme crystals were cryo-protected by paratone, and crystals of the ternary complexes were cryo-protected by transferring to their reservoir solution supplemented with 25% (v/v) glycerol. The crystals were flash-frozen in liquid propane for data collection at 100 K.

The free enzyme crystals belong to space group $P4_32_12$, with cell parameters of $a = b = 110.6$ Å and $c = 76.0$ Å. There is one molecule of CofD in the crystallographic asymmetric unit.

The two ternary complex crystals, despite being crystallized from entirely different solutions, are isomorphous to each other. They belong to space group $P3_2$ with cell parameters of $a = b = 185.3$ Å and $c = 67.8$ Å for the Fo + P_i complex, and $a = b = 186.5$ Å, $c = 67.8$ Å for the Fo + GDP complex. There are two dimers of the ternary complex in the crystallographic asymmetric unit.

Data Collection and Structure Determination—A multiple wavelength anomalous diffraction (26) data set to 3.1 Å resolution was collected on a single crystal of the CofD free enzyme at the X4A beamline of the National Synchrotron Light Source. The diffraction images were processed with the HKL package (27). The data processing statistics are summarized in Table 1.

The selenium sites were located with the program SnB (28). SOLVE/RESOLVE (29) was used for phase calculation, phase improvement, and automated model building, but only about 10% of the residues were placed. The model building was greatly facilitated by the collection of a SAD data set to 2.5 Å resolution on a crystal of the ternary complex with Fo and P_i, also at the X4A beamline of National Synchrotron Light Source. The complete atomic models, including the initiating (seleno)methionine, were built with the program XtalView (30). The structure of the ternary complex with Fo and GDP was determined by the molecular replacement method with the program COMO (31), using the structure of the ternary complex with Fo and P_i as the model.

The three structures were refined with the program CNS (32). Noncrystallographic symmetry restraint was applied for all stages of the refinement of the ternary complexes. The refinement statistics are summarized in Table 1.

CofD Activity Assays—CofD activity was measured in a coupled assay with lactylphosphate guanylyltransferase (CofC, MJ1117),[4] which produces LPPG from LP and GTP. The reaction mixture contained 50 mM TES (pH 7.5), 0.1 mM GTP, 10 mM LP, 0.4 mM Fo, 2 mM $MnCl_2$, 10 μg of CofC, and 0.4–0.6 μg of CofD. The enzyme system is inhibited by high concentrations of GMP; therefore, 2 mM phosphoenolpyruvate and 1 unit of pyruvate kinase (Sigma P-1381) were also included in the reaction mixture to regenerate GTP from GMP. The reactions were incubated at 55 °C for 30 min, after which the proteins were precipitated by addition of 1.5 volumes of methanol and removed by centrifugation. F_{420}-0 formation was determined by high pressure liquid chromatography analysis on a Varian Pursuit XRs 250 × 4.6 mm 5-μm C18 column with linear elution gradient from 95% 25 mM sodium acetate pH 6.0, 5% MeOH to 80% MeOH at 0.5 ml/min flow rate over 40 min. F_{420}-0 was detected by fluorescence (420 nm excitation and 480 nm emission).

Fluorescence Assays—All experiments were performed using a SAFAS Xenius spectrofluorometer. All spectra were corrected for buffer fluorescence. Fluorescence measurements were carried out after dilution of CofD (0.55 μM final concentration) and equilibration for 2 min in 2 ml of buffer containing 25 mM Hepes (pH 8), upon excitation at 282 nm. Binding of Fo was monitored by the variation of tryptophan-intrinsic fluorescence of CofD (between 310 and 380 nm) produced after addition of increasing concentrations of effectors. Correction for the inner filter effect was performed under the same conditions by using *N*-acetyltryptophanamide (Sigma), following a published protocol (33). Fluorescence resonance energy transfer between tryptophan residues of CofD and bound mant-nucleotide derivatives was monitored by the fluorescence emission between 310 and 500 nm. Peak integration was carried out at each ligand concentration, and the observed changes in fluorescence intensity or fluorescence resonance energy transfer were used for the calculation of ligand affinity. Data were collected at least in triplicate for each ligand, and curve fitting of the data was performed by using Graphit 4.0.10 (Erithacus Software) as described previously (34). For quenching by Fo, the Equation 1 was used,

$$F = F_{max} - (F_{max} - F_{min})\{(E + L + K_d) - [(E + L + K_d)^2 - 4EL]^{1/2}\}/2 \times E \quad \text{(Eq. 1)}$$

and for fluorescence resonance energy transfer with mant nucleotides, Equation 2 was used,

$$F = F_{min} + (F_{max} - F_{min})\{(E + L + K_d) - ((E + L + K_d)^2 - 4EL)^{1/2}\}/2 \times E \quad \text{(Eq. 2)}$$

where F is the relative fluorescence intensity; F_{max} is the maximal relative fluorescence intensity; F_{min} is the minimal relative fluorescence intensity; L is the concentration of Fo; E is the total concentration of CofD; and K_d is the dissociation constant of the enzyme-substrate complex.

BH3658 and YvcK contain no Trp residues. The binding assays monitored the fluorescence change of Fo, mant-, and trinitrophenylnucleotides in the presence of the proteins.

[4] R. H. White, unpublished results.

RESULTS AND DISCUSSION

Structure Determination—The structure of the free enzyme of *M. mazei* CofD was determined at 3.1 Å resolution by the selenomethionyl multiple wavelength anomalous diffraction method (26). However, several loops in the structure have weak electron density. The subsequent availability of a SAD data set to 2.5 Å on CofD in a ternary complex with Fo and P_i greatly facilitated the model building process.

TABLE 1
Summary of crystallographic information

Ligands	None	Fo + P_i	Fo + GDP
Maximum resolution (Å)	3.1	2.5	3.0
No. of observations	149,640	910,865	147,792
R_{merge} (%)[a]	13.4 (52.2)	8.8 (41.0)	9.3 (36.8)
$I/\sigma I$	17.9 (2.3)	15.4 (4.1)	12.1 (1.5)
Resolution range used for refinement	19.83–3.1	29.6–2.5	28.7–3.0
No. of reflections[b]	13,019	164,849	39,441
Completeness (%)	80 (44)	92 (76)	75 (46)
R factor (%)[c]	23.9 (37.3)	19.5 (24.5)	20.6 (33.4)
Free *R* factor (%)	27.7 (38.8)	23.1 (28.9)	23.7 (33.7)
Root mean square deviation in bond lengths (Å)	0.009	0.007	0.009
Root mean square deviation in bond angles (°)	1.2	1.2	1.2
PDB accession code	3CGW	3C3D	3C3E

[a] $R_{merge} = \Sigma_h\Sigma_i|I_{hi} - \langle I_h\rangle|/\Sigma_h\Sigma_i I_{hi}$. The numbers in parentheses are for the highest resolution shell.
[b] The number for the selenomethionyl protein includes both Friedel pairs.
[c] $R = \Sigma_h|F_h^o - F_h^c|/\Sigma_h F_h^o$.

The current atomic models of the three structures of CofD, free enzyme and the ternary complexes with Fo + P_i and Fo + GDP, contain residues 1–303 for all the protein molecules in the asymmetric unit, together with a few non-native residues from the C-terminal His tag (LEHHHHHH). All the structures have low *R* values and excellent geometry at their respective resolution (Table 1). The majority (90%) of the residues are located in the most favored region of the Ramachandran plot, and no residues are located in the disallowed region (data not shown).

Crystals of the two ternary complexes contain four CofD monomers in the crystallographic asymmetric unit. The structures of the monomers in each crystal are essentially identical. There are, however, recognizable differences between the conformations of the two ternary complexes in the active site region, and these will be described later.

Overall Structure of CofD—The structure of CofD monomer contains 12 β-strands (named β1 through β12) and 13 α-helices (αA through αM) (Fig. 1*B* and Fig. 2). Seven of the β-strands (β1, β2, β5, β9–β12) form a Rossmann fold, but there are three large inserted segments between consecutive β-strands in this fold that decorate the C-terminal end of the central parallel β-sheet. The first insert, with about 75 residues and located in the connection between neighboring strands β2 and β5, contains a small β-hairpin (β3 and β4), 3 helices, and 2 long loops (Fig. 2). The second insert (with 40 residues and located between strands β5 and β9) is composed of a three-stranded anti-parallel β-sheet (β6, β7, and β8) and a helix. Finally, the third insert (with 20 residues and between β10 and β11) contains a helix and a long loop. All three inserts have important roles in forming the dimer interface and/or the active site of CofD (Fig. 2 and see below).

Earlier gel filtration data showed that CofD is dimeric in solution (20). The structures show a highly elongated dimer (Fig. 2), and this dimeric association is observed in all three structures, in two different crystal forms. The dimer interface is formed by the first insert to the

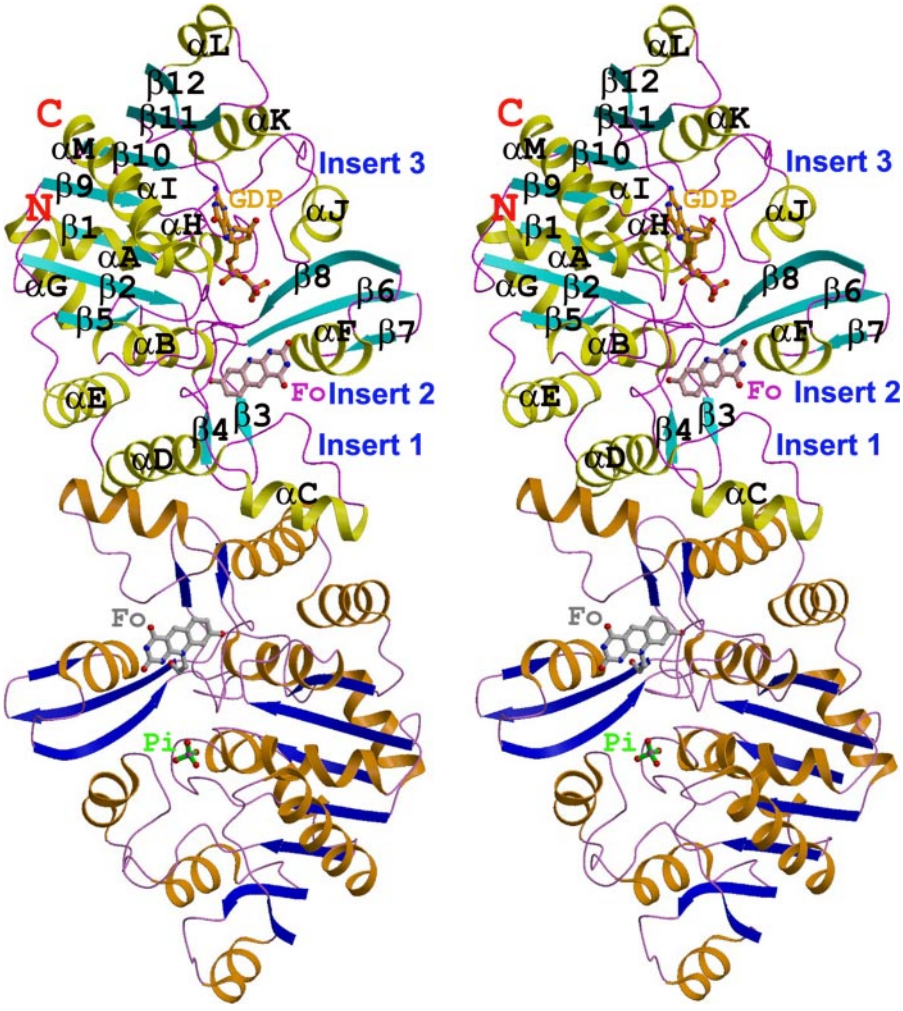

FIGURE 2. **Structure of the CofD dimer.** Schematic representation of the structure of the CofD dimer. One monomer shows the ternary complex with Fo (*pink*) and GDP (*brown*), and the other with Fo (*gray*) and P_i (*green*). Unless otherwise noted, all structural figures were produced with Molscript (39) and rendered with Raster3D (40).

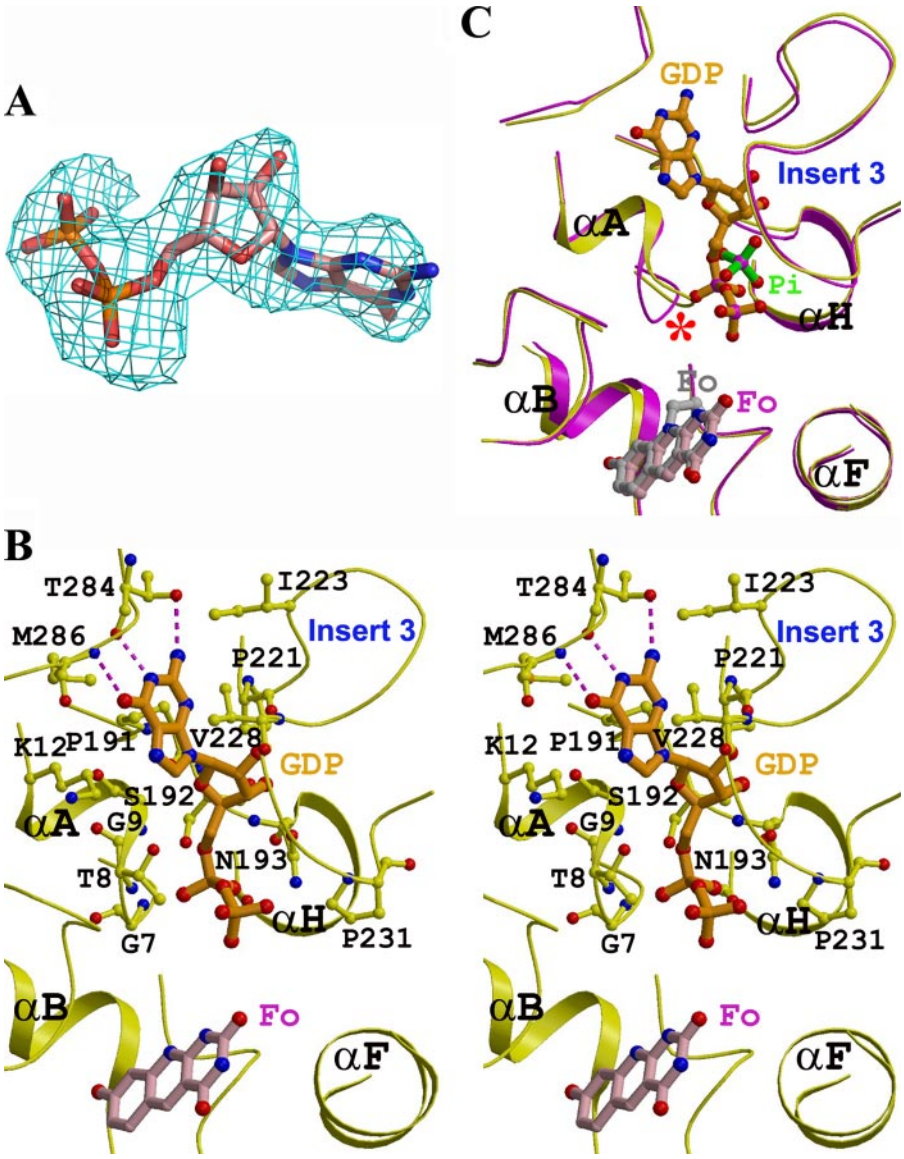

FIGURE 3. **Binding mode of GDP.** *A,* final $2F_o - F_c$ electron density for GDP molecule, contoured at 1σ. Produced with PyMol (41). *B,* schematic drawing of the interactions between GDP and CofD. *C,* overlay of the structure of the ternary complex with Fo and GDP (in *purple*) and that with Fo and P_i (in *yellow*). The P_i is located about 1 Å from the α-phosphate of GDP, but there is a large difference in the conformation of the glycine-rich loop, indicated with the *red star*.

Rossmann fold and buries about 950 Å^2 of the surface area of each monomer (Fig. 2). The β3–β4 hairpin and the αC, αD helices in one monomer are in contact with their equivalents in the other monomer in this interface. As some of these residues are also involved in binding the deazariboflavin of Fo (Fig. 2, see below), it is possible that dimerization is important for stabilizing the conformation of this insert for substrate binding. Currently, it is not known whether dimerization is required for the catalytic activity of CofD.

Binding Mode of GDP—The active site of CofD is located at the C-terminal end of the central β-sheet of the Rossmann fold, with contributions from the three major inserts to this fold (Fig. 2). To define the binding modes of the substrates of this enzyme, we included Fo and GTP in one set of crystallization experiments and succeeded in obtaining the structure of a ternary complex with Fo and GDP. By including Fo and LP in the crystallization solution, we were able to determine the structure of another ternary complex, with Fo and P_i. Although LPPG is the actual substrate of the enzyme, it is not stable in solution, with a half-life of about 2 h at room temperature (20, 21). As LPPG is produced by the condensation of LP and GTP (21), we included these two compounds instead in our crystallization experiments. Although 5 mM GTP was used in the crystallization solution, only the electron density for GDP is observed in the four CofD molecules based on the crystallographic analysis (Fig. 3*A*). These crystals took 2 weeks to grow, and it is possible that GTP has been hydrolyzed during this time, or the γ-phosphate is disordered in structure because of a lack of strong interactions with the enzyme (see below).

The GDP molecule is bound between the central Rossmann fold core and the third insert of CofD (Fig. 2). Many of the residues that interact with GDP are conserved among the F_{420}-producing CofD homologs (Fig. 1*B*), confirming their functional importance. One face of the guanine base lies against the side chains of Lys-12, Pro-191, and Met-286, and the other face has van der Waals interactions with the side chains of Ile-223 and Val-228 in the long loop in the third insert (Fig. 3*B*). The N-1 and O-6 atoms of the guanine base are recognized by hydrogen-bonding interactions with the main chain atoms of residues Thr-284 and Met-286 (Fig. 3*B*), in the loop connecting the last strand (β12) and last helix (αM) of the Rossmann fold (Fig. 2). The 2-amino group of guanine is hydrogen-bonded to the main chain carbonyl of Ser-220 and the side chain hydroxyl of Thr-284, which in turn interacts with the side chain of Asp-262, although Thr-284 is not conserved among the CofD homologs (Fig. 1*B*). This pattern of specific recognition of the guanine base is entirely different from that observed in G proteins. Nonetheless, the structure shows that the adenine base cannot maintain the same hydrogen-bonding interactions with the enzyme, consistent with kinetic data showing that the K_m value for lactyl(2)-diphospho-(5′)adenosine is about 30-fold higher than LPPG (20).

The α-phosphate of GDP has many hydrogen-bonding interactions with the enzyme, primarily with the glycine-rich loop in the connection between the first strand (β1) and first

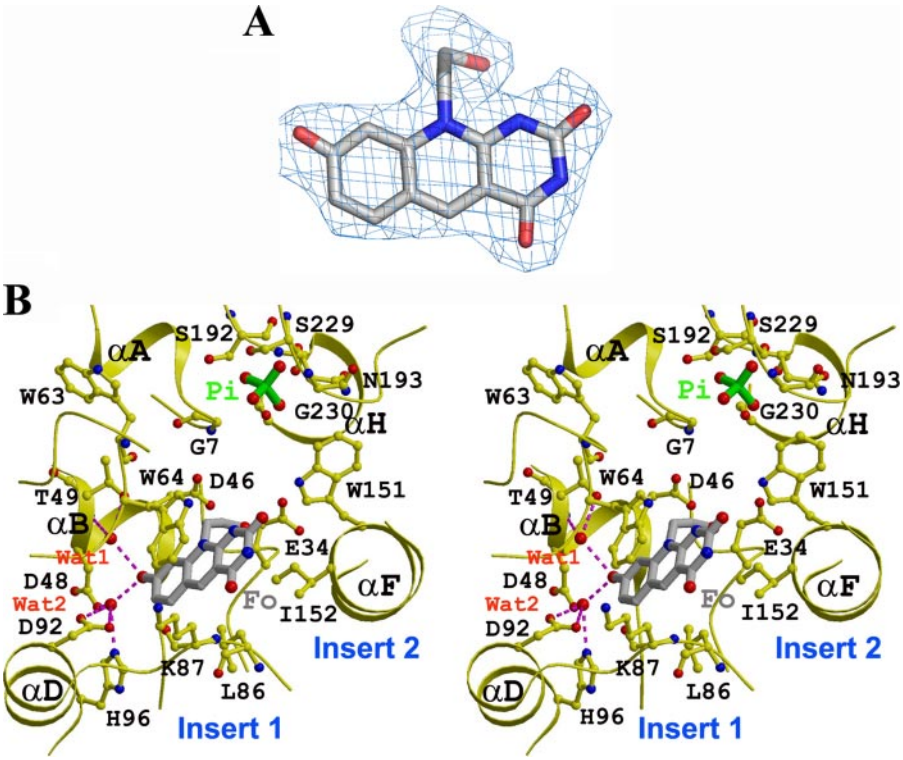

FIGURE 4. **Binding mode of Fo.** *A*, final $2F_o - F_c$ electron density for Fo, contoured at 1σ. Produced with PyMol (41). *B*, schematic drawing of the interactions between Fo and CofD.

helix (αA) of the Rossmann fold (Fig. 1*B*), as well as with the side chains of Ser-192, Asn-193, and Ser-197 (Fig. 3*B*). These three residues are part of the PSNP*XX*SI motif that is conserved among the F_{420}-producing CofD homologs (Fig. 1*B*) (20). In comparison, the β-phosphate of GDP does not appear to be specifically recognized by the enzyme in this structure (Fig. 3*B*). The 2′-hydroxyl of the ribose is hydrogen-bonded to the main chain atoms of Pro-221 and Val-228, whereas the 3′-hydroxyl interacts with the side chain of Asn-193 (Fig. 3*B*).

Crystallographic analysis on the crystal grown in the presence of Fo and LP revealed the presence of a phosphate group in the electron density. The lactyl group is either disordered or it has been hydrolyzed during crystallization. Another possibility is that our sample of LP contained some free phosphates. This phosphate is located about 1 Å away from the α-phosphate group of GDP, and the glycine-rich loop has a significant conformational difference in this complex (Fig. 3*C*).

Binding Mode of Fo—The deazariboflavin ring of Fo has well defined electron density in the ternary complex with Fo and P_i (Fig. 4*A*). This group has a planar conformation in our structure, in contrast to a butterfly conformation as observed in F_{420}-dependent dehydrogenases (35–37). In addition, the mode of recognition of Fo is highly different between CofD and F_{420}-dependent dehydrogenases (35–37). The ribityl group of Fo has much weaker electron density (Fig. 4*A*). A satisfactory model could not be built for it based on the crystallographic analysis, and most of this group has been omitted from the current atomic model. The electron density for Fo in the ternary complex with Fo and GDP is weaker but shows similar interactions with the enzyme as those observed in the Fo + P_i ternary complex.

The binding site for Fo is formed by residues in the first and second inserts of CofD (Fig. 2). Fo has both hydrogen-bonding interactions with conserved residues in CofD, mediated through two water molecules, and van der Waals interactions with the enzyme. One face of the deazariboflavin group is π-stacked against the side chain of Trp-64 (in the first insert), and the other face has van der Waals interactions with the side chains of Ile-43, Leu-86, and Ile-152 (Fig. 4*B*). The 8-hydroxyl group on the phenyl ring interacts through two water molecules with several residues in CofD, including the side chains of Asp-48 and Asp-92, from the two conserved D*X*D motifs in F_{420}-producing CofD homologs (Fig. 1*B*) (20). Such interactions help the enzyme distinguish between the 7,8-didemethyl-8-hydroxy-5-deazariboflavin in Fo and regular riboflavin, as the latter is not a substrate for CofD (20). Moreover, our structural analysis shows that introducing a methyl group at the 7 position of Fo will cause steric clashes with the side chain of Asp-92, which should also help the discrimination against regular riboflavin by CofD. The carbonyl groups on the pyrimidine ring of Fo are exposed to the solvent (Fig. 4*B*).

Large Conformational Changes upon Substrate Binding—A comparison of the structures of the free enzyme of CofD and the two ternary complexes shows that there are large conformational changes in the enzyme upon substrate binding, primarily in the second and third inserts (Fig. 5*A*). The largest difference is seen for the loop connecting strand β10 and helix αJ, in the third insert. This loop is located away from the GDP molecule in the free enzyme structure, creating an open form of the active site (Fig. 5*B*). In the ternary complex, this loop moves into the active site (Fig. 5*A*), creating a closed form of the enzyme (Fig. 5*C*) and providing several side chains that are important for interactions with the guanine base and the ribose. Interestingly, the structure of the Fo + P_i ternary complex has a similar conformation for this loop (Fig. 5*A*), suggesting that a single phosphate group may be sufficient to induce active site closure. However, the glycine-rich loop assumes a different conformation in the ternary complex with P_i, as discussed earlier (Fig. 3*C*).

Additional differences are seen in the position of helix αF, in the second insert. It moves closer to the active site to establish interactions with the Fo substrate (Fig. 5*A*). A conformational change is also observed for helix αL, in the connection between strands β11 and β12 in the Rossmann fold core, but this helix is located far from the active site.

Substrate Binding Assays Confirm the Structural Observations—Our structures show that CofD contains three Trp residues, Trp-63, Trp-64, and Trp-151, in the active site region (Fig. 4*B*), which allowed us to develop assays that can monitor ligand binding by either intrinsic Trp fluorescence or by fluorescence resonance energy transfer. In particular, Trp-64 is π-stacked with the deazariboflavin moiety of Fo (Fig. 4*B*) and therefore is the most likely to change its fluorescence upon Fo binding. Indeed, in the absence of Fo and with the excitation wavelength set at 282 nm, CofD exhibited a fluorescence emission spectrum characteristic of a rather accessible tryptophan residue (λ_{max} emission around 340 nm; data not shown). A progressive increase of the Fo concentration from 1.25 to 20 μM quenched the fluorescence intensity of CofD (Fig. 6*A*), from which the apparent K_d for the binding of Fo to CofD can be estimated to be 1.38 $\pm$ 0.07 μM.

We also attempted to monitor the binding of GDP to CofD by intrinsic Trp fluorescence. The addition of GDP produced a weak specific change of CofD fluorescence, but the curve fitting of the data did not allow us to determine a precise apparent K_d value (data not shown). We then used the fluorescent nucleotide mant-GDP in the binding assay. Addition of increasing amounts of mant-GDP produced a quenching of the fluorescence emission spectrum of CofD tryptophan residues (Fig. 6*B*). Simultaneously, a new fluorescence signal progressively developed, centered at ~430 nm, which is related to fluorescence resonance energy transfer between the tryptophan residues and the mant group. The fluorescence resonance energy transfer data allowed us to estimate an apparent K_d of 1.59 $\pm$ 0.14 μM for the CofD-mant-GDP complex (Fig. 6*C*). A K_d value of 4.6 $\pm$ 0.6 μM was determined for the CofD-mant-GTP complex using similar experiments. No binding of mant-ADP or

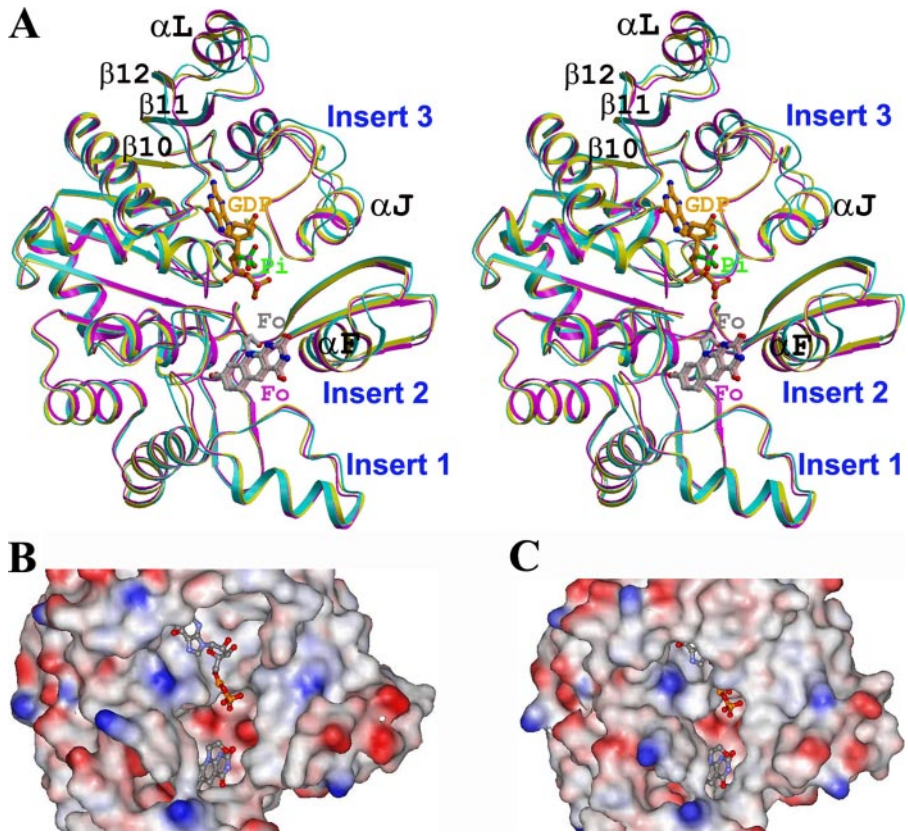

FIGURE 5. **Induced fit behavior of CofD.** *A,* overlay of the structures of the free enzyme of CofD (in *cyan*), the ternary complex with Fo and GDP (in *purple*), and the ternary complex with Fo and P_i (in *yellow*). *B,* molecular surface of CofD in the active site region in the free enzyme in an *open state*. The structures of GDP and Fo are shown for reference. *C,* molecular surface of CofD in the active site region in the ternary complex with Fo and GDP in a *closed state*.

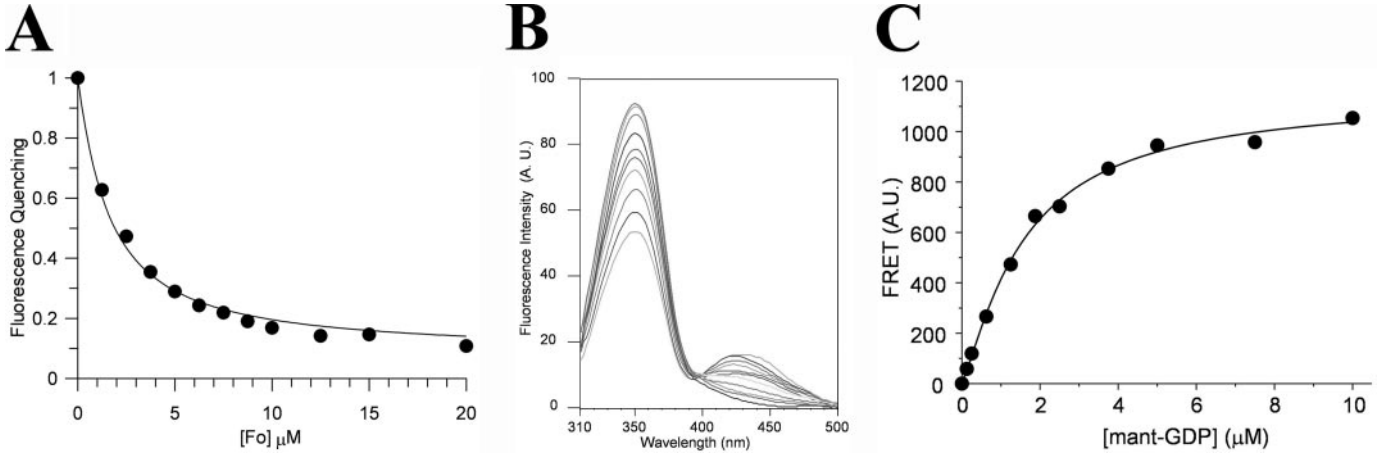

FIGURE 6. **Fluorescence binding assays for CofD.** *A,* fluorescence quenching of CofD as a function of Fo concentration. *B,* emission spectrum of CofD recorded at different concentrations of mant-GDP, with excitation at 282 nm. *C,* fluorescence resonance energy transfer, taken as the increase in fluorescence between 400 and 500 nm, is plotted against the concentration of mant-GDP.

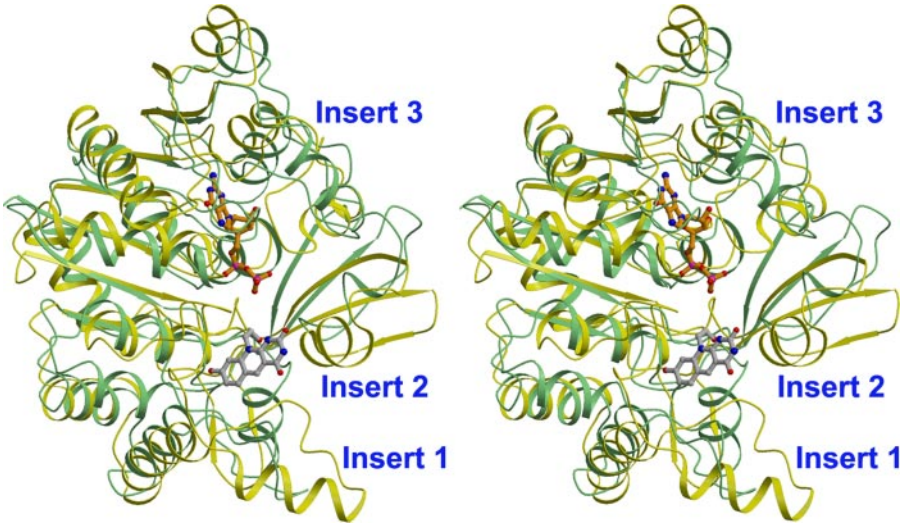

FIGURE 7. **CofD homologs from non-F_{420}-producing organisms have large differences in structure.** Overlay of the structure of CofD from *M. mazei* (in *yellow*) and BH3568 protein from *B. halodurans* (in *green*).

mant-ATP was observed at up to 75 μM concentration, consistent with our structural observations (Fig. 3*B*) and earlier kinetic data (20).

Implication for the Biosynthesis of F_{420}—Our structural studies have defined the binding modes of GDP and the deazariboflavin portion of Fo, as well as located the active site region of the CofD enzyme. The structural analysis suggests that the GMP portion of the LPPG substrate of CofD is likely to have a similar binding mode. Although the exact binding modes of the LP portion of LPPG and the ribityl group of Fo are not defined from these structures, a careful examination of the active site region of CofD provides significant insights into the mechanism of catalysis by this enzyme. The reaction is known to require Mg^{2+} ion(s), which are likely coordinated by the phosphate groups and the carboxylate group of the LPPG substrate (20). Our structures show that two acidic side chains in the active site region, Glu-34 and Asp-46 (Fig. 4*B*), could also be ligands to the Mg^{2+} ions. Asp-46 is part of the first conserved D*X*D motif, and Glu-34 is conserved as an acidic residue among F_{420}-producing CofD homologs (Fig. 1*B*). To confirm the functional importance of the Asp-46 residue, we mutated its equivalent residue in the *M. jannaschii* CofD homolog to alanine and observed a 40-fold loss in the specific activity of the enzyme.

It is likely that the conformation of the LP and ribityl groups of the substrates are only ordered in the presence of Mg^{2+} ions, which brings the terminal hydroxyl group of Fo into proximity of the β-phosphate of LPPG to initiate the reaction. Our structural analysis suggests that the general base that extracts the proton from this hydroxyl group is likely a hydroxide ligated to the Mg^{2+} ions or possibly one of the terminal oxygens on the LPPG substrate, as the side chains of acidic residues in the active site region are likely too far from this hydroxyl group.

Our structures also show that only the α-phosphate of GDP is tightly anchored by its interactions with the enzyme (Fig. 3*B*). Therefore, other compounds that contain di- or triphosphate groups could become substrates to the enzyme, by overlapping with this phosphate. This explains the observations that CofD can catalyze the phosphorylation of Fo by using GTP, GDP, PP_i, and PPP_i as substrates (20).

Large Structural Differences in Non-F_{420}-producing CofD Homologs—CofD belongs to a large family of proteins that have been found in both F_{420}-producing and non-F_{420}-producing organisms (Fig. 1*B*). The crystal structures of several non-F_{420}-producing CofD homologs have been determined, including the BH3568 protein from *Bacillus halodurans* (PDB entries 2HZB and 2O2Z, 24% sequence identity to *M. mazei* CofD), the LP_0780 protein from *Lactobacillus plantarum* (PDB entry 2P0Y, 21% identity), and the SE_0549 protein from *Staphylococcus epidermidis* (PDB entry 2PPV, 18% identity). The central Rossmann fold of these structures is similar. The root mean square distance between equivalent C-α atoms of the structures of CofD and BH3568 is 3.3 Å (Z score of 22.6 from the program Dali) (38).

However, our detailed sequence and structural and functional analyses suggest that the non-F_{420}-producing CofD homologs may use completely different substrates and catalyze a different biochemical reaction compared with CofD. Evidence in support of this conclusion includes the following. 1) Many of the conserved sequence motifs in the F_{420}-producing homologs (for example the two D*X*D motifs and the PSNP*XX*SI motif) are not conserved in the non-F_{420}-producing homologs (Fig. 1*B*). Our structural data on CofD show that these sequence motifs are important for GDP binding and/or catalysis. 2) There are large differences in the conformation of the inserts, and as a result there are large differences in the active site of the non-F_{420}-producing CofD homologs (Fig. 7). The structure of the first insert of BH3568 is entirely different from that in CofD (Fig. 7). As this insert is involved in dimerization, the BH3568 dimer has a significantly different organization relative to the CofD dimer as well. This difference is also expected to disrupt the Fo-binding site, and the Trp-64 residue in CofD does not have an equivalent in BH3568. The second insert in BH3568 has a similar fold but is positioned differently compared with CofD (Fig. 7). The third insert is much shorter in BH3568, and this protein contains a new insert, between strand β11 and helix αL (Fig. 7). In addition, the loop connecting strand β12 and helix αM is longer (by 4 residues, Fig. 1) and assumes a different conformation in BH3568. As residues in this loop are involved in recognizing the guanine base in CofD (Fig. 3*B*), the structural difference suggests that the non-F_{420}-producing homologs probably cannot bind GDP. 3) We did not detect any binding of Fo, GDP, ADP, GTP, or ATP to the BH3568 protein or its close homolog YvcK from *B. subtilis* in our assays (22). The binding data provide strong experimental support for the observations based on the sequence and structural analyses. Further studies are needed to characterize the exact biochemical and biological functions of the non-F_{420}-producing CofD homologs. On the other hand, the glycine-rich loop is conserved in these proteins (Fig. 1*B*), suggesting that the substrate(s) is likely to contain a

phosphate group. In fact, one of the BH3568 structures (PDB entry 2O2Z) contains a sulfate at this position.

In summary, our structural and biochemical studies have provided significant molecular insights into one of the steps in the biosynthesis of the F_{420} coenzyme. The structural analysis also indicates the presence of a structurally homologous family of enzymes that likely catalyze a different biochemical reaction.

Acknowledgments—We thank Randy Abramowitz and John Schwanof for setting up the X4A beamline, and Sergey M. Vorobiev and Jordi Benach for assistance in crystal screening and data collection.

REFERENCES

1. Graham, D. E., and White, R. H. (2002) *Nat. Prod. Rep.* **19,** 133–147
2. Lessner, D. J., Li, L., Li, Q., Rejtar, T., Andreev, V. P., Reichlen, M., Hill, K., Moran, J. J., Karger, B. L., and Ferry, J. G. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 17921–17926
3. DiMarco, A. A., Bobik, T. A., and Wolfe, R. S. (1990) *Annu. Rev. Biochem.* **59,** 355–394
4. Johnson, E. F., and Mukhopadhyay, B. (2005) *J. Biol. Chem.* **280,** 38776–38786
5. McCormick, J. R. D., and Morton, G. O. (1982) *J. Am. Chem. Soc.* **104,** 4014–4015
6. Rhodes, P. M., Winskill, N., Friend, E. J., and Warren, M. (1981) *J. Gen. Microbiol.* **124,** 329–338
7. Coats, J. H., Li, G. P., Kuo, M. S., and Yurck, D. A. (1989) *J. Antibiot.* (*Tokyo*) **42,** 472–474
8. Ikeno, S., Aoki, D., Hamada, M., Hori, M., and Tsuchiya, K. S. (2006) *J. Antibiot.* (*Tokyo*) **59,** 18–28
9. Peschke, U., Schmidt, H., Zhang, H. Z., and Piepersberg, W. (1995) *Mol. Microbiol.* **16,** 1137–1156
10. Eker, A. P., Hessles, J. K. C., and van de Velde, J. (1988) *Biochemistry* **27,** 1758–1765
11. Eker, A. P., Kooiman, P., Hessles, J. K. C., and Yasui, A. (1990) *J. Biol. Chem.* **265,** 8009–8015
12. Manjunatha, U. H., Boshoff, H. I. M., Dowd, C. S., Zhang, L., Albert, T. J., Norton, J. E., Daniels, L., Dick, T., Pang, S. S., and Barry, C. E., III (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 431–436
13. Boshoff, H. I. M., and Barry, C. E., III (2005) *Nat. Rev. Microbiol.* **3,** 70–80
14. Stover, C. K., Warrener, P., VanDevanter, D. R., Sherman, D. R., Arain, T. M., Langhorne, M. H., Anderson, S. W., Towell, J. A., Yuan, Y., McMurray, D. N., Kreiswirth, B. N., Barry, C. E., and Baker, W. R. (2000) *Nature* **405,** 962–966
15. Bair, T. B., Isabelle, D. W., and Daniels, L. (2001) *Arch. Microbiol.* **176,** 37–43
16. Li, H., Graupner, M., Xu, H., and White, R. H. (2003) *Biochemistry* **42,** 9771–9778
17. Li, H., Xu, H., Graham, D. E., and White, R. H. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 9785–9790
18. Graham, D. E., Xu, H., and White, R. H. (2003) *Arch. Microbiol.* **180,** 455–464
19. Grochowski, L. L., Xu, H., and White, R. H. (2006) *J. Bacteriol.* **188,** 2836–2844
20. Graupner, M., Xu, H., and White, R. H. (2002) *Biochemistry* **41,** 3754–3761
21. Graupner, M., and White, R. H. (2001) *Biochemistry* **40,** 10859–10872
22. Gorke, B., Foulquier, E., and Galinier, A. (2005) *Microbiology* **151,** 3777–3791
23. Acton, T. B., Gunsalus, K., Xiao, R., Ma, L., Aramini, J., Baron, M. C., Chiang, Y., Clement, T., Cooper, B., Denissova, N., Douglas, S., Everett, J. K., Palacios, D., Paranji, R. H., Shastry, R., Wu, M., Ho, C.-H., Shih, L., Swapna, G. V. T., Wilson, M., Gerstein, M., Inouye, M., Hunt, J. F., and Montelione, G. T. (2005) *Methods Enzymol.* **394,** 210–243
24. Jansson, M., Li, Y.-C., Jendeberg, L., Anderson, S., Montelione, G. T., and Nilsson, B. (1996) *J. Biomol. NMR* **7,** 131–141
25. Doublie, S., Kapp, U., Aberg, A., Brown, K., Strub, K., and Cusack, S. (1996) *FEBS Lett.* **384,** 219–221
26. Hendrickson, W. A. (1991) *Science* **254,** 51–58
27. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
28. Weeks, C. M., and Miller, R. (1999) *J. Appl. Crystallogr.* **32,** 120–124
29. Terwilliger, T. C. (2003) *Methods Enzymol.* **374,** 22–37
30. McRee, D. E. (1999) *J. Struct. Biol.* **125,** 156–165
31. Jogl, G., Tao, X., Xu, Y., and Tong, L. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57,** 1127–1134
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
33. Pompeo, F., Luciano, J., and Galinier, A. (2007) *J. Bacteriol.* **189,** 1154–1157
34. Divita, G., Goody, R. S., Gautheron, D. C., and di Pietro, A. (1993) *J. Biol. Chem.* **268,** 13178–13186
35. Aufhammer, S. W., Warkentin, E., Ermler, U., Hagemeier, C. H., Thauer, R. K., and Shima, S. (2006) *Protein Sci.* **14,** 1840–1849
36. Warkentin, E., Mamat, B., Sordel-Klippert, M., Wicke, M., Thauer, R. K., Iwata, M., Iwata, S., Ermler, U., and Shima, S. (2001) *EMBO J.* **20,** 6561–6569
37. Aufhammer, S. W., Warkentin, E., Berk, H., Shima, S., Thauer, R. K., and Ermler, U. (2004) *Structure* (*Lond.*) **12,** 361–370
38. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
39. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
40. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524
41. DeLano, W. L. (2002) *The Pymol Manual,* DeLano Scientific, San Carlos, CA